2015

Experience in using MTurk for Network Measureme

Huz, Gokay

þÿ C 2 B (1) D 1 5 , A u g u s t 1 7 , 2 0 1 5 , L o n d o n , U n i t e d K i n g d o m
http://hdl.handle.net/10945/46376

Experience in using Mechanical Turk for Network Measurement

Gokay Huz (NPS), **Steven Bauer** (MIT),
kc claffy (CAIDA), Robert Beverly (NPS)

ACM SIGCOMM C2B(I)D Workshop
August 21, 2015

The Problem: Internet Measurement

- Internet measurement is hard:
  - Network is large, complex, and dynamic
  - Not designed to be measured
  - Inherent security and privacy concerns
- Key concern: passive or active vantage point
  - Can strongly influence resulting data / inference
  - But, researchers typically have access to few VPs

The Problem: Internet Measurement

- Internet measurement is hard:
  - Network is large, complex, and dynamic
  - Not designed to be measured
  - Inherent security and privacy concerns
- Key concern: passive or active vantage point
  - Can strongly influence resulting data / inference
  - But, researchers typically have access to few VPs

Obtaining representative and sound Internet measurement datasets is challenging

Vantage Points (VPs)

- For network research, more VPs useful:
  - During exploratory phase, collect data from interesting/varied locations
  - For large, Internet-wide studies to collect as much data as possible
  - When attempting to validate or correlate results obtained from other methods/tools
- ***Crowdsourcing*** is a means to obtain more VPs

Crowdsourced Network Measurement

- Prior work has leveraged crowdsourcing for network measurement e.g., [Choffnes10, Bischof11]
- Our work examines Amazon's Mechanical Turk (MTurk):
  - Traditionally used for Human Intelligence Tasks (HITs), not measurement
  - E.g., surveys, annotation/labeling, psychology
  - Instead, we create network measurement HITs

MTurk Measurement HITs

- We experimented with three HITs:
  1. Broadband speed testing from select VPs (exploratory)
  2. Testing a network security property (large scale collection)
  3. IPv6 adoption (validation)
- We obtained IRB exemption from our institution
- While our results are mostly anecdotal, we hope these initial experiences are valuable for future measurement research

Broadband Speed Test

- Residential broadband speed testing has seen significant recent work (Speedtest, NDT) and research [Bauer10]
- We wished to understand the performance of new, very high speed broadband links (100Mbps – 1Gbps), e.g., Google Fiber
- However, we had no measurements and no access to any VPs on these networks
- Exploratory solution: targeted HITs

Spoofer Testing

- Understanding the susceptibility of the Internet to spoofed-source attacks is critical
- Prior work [Beverly09] solicited volunteers to run the Spoofer testing tool
- More measurements from more VPs would increase the soundness of the results
- We created a HIT that required workers to run the Spoofer tester and report results

(Not) Testing Spoofing

- Prior work [Christin12] similarly required workers to download and run an executable
- Within hours, our HIT was reported and removed:
  - Christin too was removed, but then unblocked
  - We had no such luck when contacting Amazon

(Not) Testing Spoofing

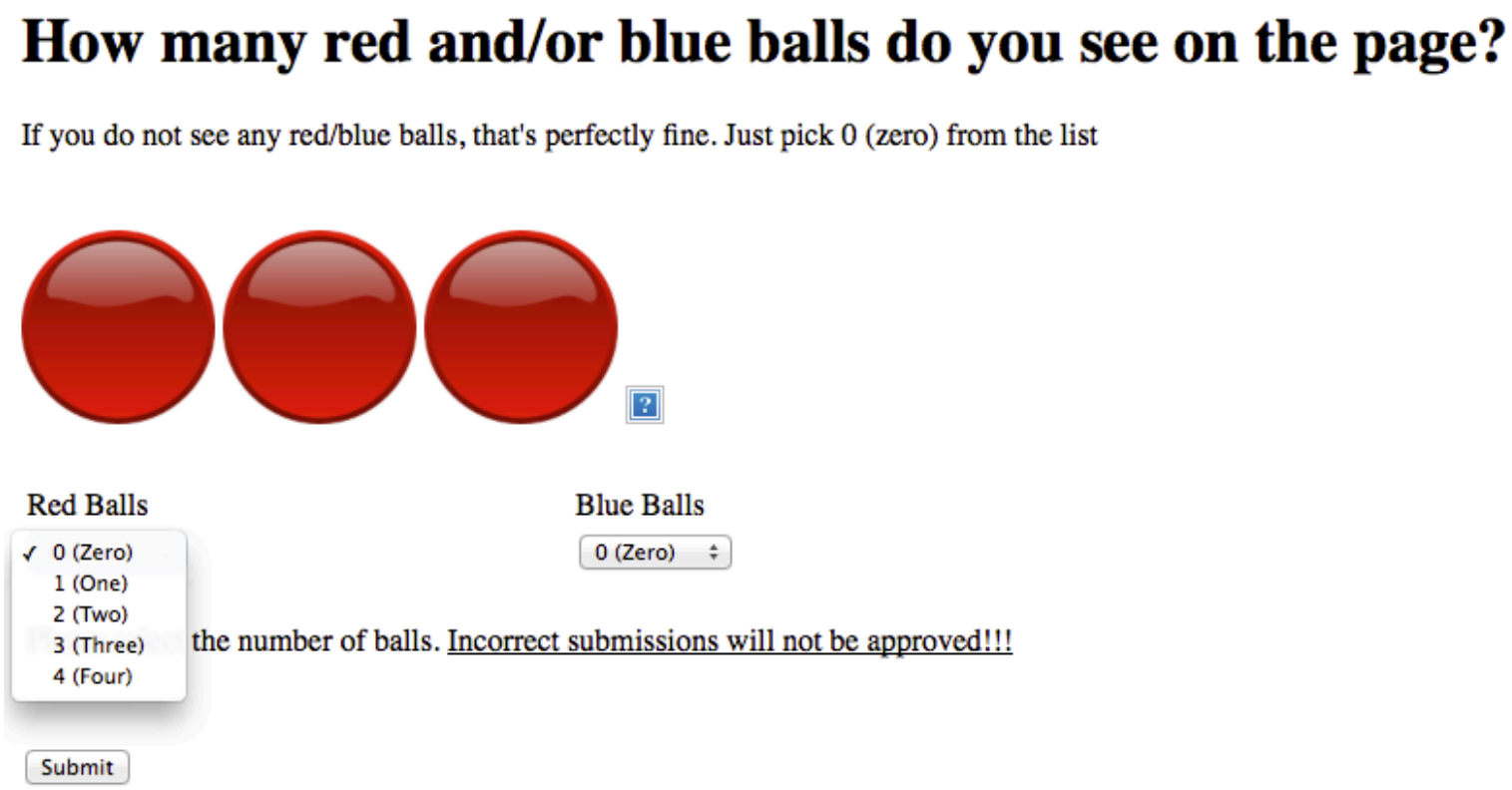
- ToS:
  - Forbade "*HITs that require workers to download software*"
  - Since relaxed to: "*HITs that require workers to download software that contains any malware, spyware, viruses, or other harmful code*"
- Suggests that researchers may be more successful in enlisting workers to run software HITs in the future

IPv6 Adoption

- Prior work has investigated IPv6 adoption
- Our goal:
  - *Embed* measurement within a HIT
  - Where the actual work of the HIT is incidental to the measurement we're collecting
  - Analyze the distribution of networks running our HIT
  - Compare to other IPv6 adoption results

IPv6 Measurement HIT: Ball Counting

- User is given a random number of red and blue balls. For example:



IPv6 Measurement HIT: Ball Counting

- We host the URL for the balls
  - Red balls availabile via IPv4 only
  - Blue balls available via IPv6 only
  - The blue ball URL is a function of the user's IPv4 address
  - `http://ipv6.example/img.php?1.2.3.4`
- For a user with IPv6 connectivity, our web server logs:

```
2001:dead::beef:cafe - - [11/Mar/
2014:01:17:36] "GET /img.php?1.2.3.4 HTTP/
1.1" 200 37977 "http://ipv4.example/?
assignmentId=XXXXXX
&hitId=YYYYYY&workerId=ZZZZZZ"
```

- Allowing us to match the client's IPv4 and IPv6 addresses

Results

- Next, we present result highlights
- These initial results intended to describe our experience and generate discussion
- See paper for more details…

A New Form of Bias: Worker Networks

- Among IPv4 workers:
  - 322 (60.8%) of workers geolocate to US
  - 148 (27.9%) geolocate to India
  - 58 (11.3%) geolocate elsewhere
- Among 37 IPv6 capable clients:
  - 20 used Teredo or 6to4
  - 17 geolocate to US
- Suggests:
  - Homogeneity in countries and networks from which researchers can expect measurement results
  - Difficulty in obtaining diverse VPs

Previewing HITs

- We obtain measurement results when workers preview our HIT:
  - Generates HTTP fetches required for IPv6 adoption inference
  - Without completing HIT or providing compensation
  - Unintentional; interesting incidental finding

Over constrained HITs

- Further, we find that it is possible to create over-constrained HITs:
  - E.g., worker must be in both Japan and US
  - Workers cannot accept or perform HIT
  - Workers therefore cannot be compensated
  - However, workers still preview HIT, generating measurements
- Suggest that Amazon fix to prevent

Limitations and Future Work

- While we find a general lack of geographic and network diversity, in the future we need to control for time-of-day and native language
- We plan to attempt our spoofer measurement again, after discussing with Amazon
- We are currently experimenting with other measurements using MTurk, including IP geolocation and DNSSEC validation

Thanks!

- Questions?